Int. J. Appl. Math. and Comp. Sci., *1999*, *Vol.9*, *No.1*, *25–52*

ADAPTIVE CONTROL THROUGH ON-LINE OPTIMIZATION. THE MPCI PARADIGM AND VARIANTS

MICHAEL NIKOLAOU*, S. ALPER EKER**

A new variant of Model Predictive Control and Identification (MPCI) is proposed. The on-line objective is not to minimize the sum of square errors, but to maximize on-line the sum of the lower bounds on the minimum eigenvalues of the information matrices over finite horizons. In that way, inputs to the controlled process are allowed to excite the process highly enough to generate as much modelling information as possible, while the process goes off-spec as little as possible. Constraints can be loosened or tightened according to the need for identification. The effectiveness of the proposed new methodology is illustrated through a number of simulations.

Keywords: Adaptive control, model predictive control, closed-loop identification, persistent excitation, linear matrix inequalities, semidefinite programming.

1. Introduction

Model Predictive Control (MPC) is a class of computer control algorithms that explicitly use a process model to predict future plant outputs and compute an appropriate control action through on-line optimization of a cost objective over a future horizon, subject to various constraints. Performance of MPC could become unacceptable due to a very inaccurate model, thus requiring a more accurate model. Such a model frequently has to be developed while the process is kept under MPC. This task is an instance of closed-loop identification and adaptive control. The difficulty of closed-loop identification is that the input of the process to be identified is not directly selected by the designer but ultimately by the feedback controller. According to a standard closed-loop identification approach, the designer may indirectly influence the process input by exciting the feedback loop through external dithering signals (Anderson and Johnson, 1982). The difficulty of this approach lies in the fact that the effect of the external dithering signals on the process input depends on the entire closed loop, which includes the process to be identified. Consequently, guaranteeing that the process input will indeed be good for identification is not always easy and

* Dept. of Chemical Engineering, University of Houston, Houston, TX 77204–4792, U.S.A., e-mail: nikolaou@uh.edu.

** Dept. of Chemical Engineering, Texas A&M University, College Station, TX 77843–3122, U.S.A., e-mail: ekera@jetson.uh.edu.

hinges on assumptions such as knowing exactly the order of the process model, or assuming that input constraints are not an issue.

Of course, external excitation is not necessary for closed-loop identification. The self-tuning regulator (STR) with relay (Åström and Wittenmark, 1989, p.332), for example, is a simple method that uses a non-invertible controller (relay) to get around the trap of identifying the plant as minus the inverse of the controller (Nikolaou, 1998). Similarly, the Ziegler-Nichols ultimate-cycling tuning method (Ziegler and Nichols, 1942) uses a time-varying feedback gain to identify a linear time-invariant process, an idea that has strong theoretical justification (Ljung, 1987, p.366). No inputs external to the loop are used in either case. Example 2 below shows that closed-loop identification is possible even for a linear time invariant process with a proportional controller without any external excitation.

A key requirement for identification, either in open or closed-loop mode, is that *data relevant to the structure of the model to be identified demonstrate persistent excitation (PE).* By ensuring that the corresponding information matrix is invertible, PE ensures that the least-squares problem corresponding to parameter identification has a unique solution (see Section 2). When closed-loop identification is coupled with controller adaptation, lack of PE may result in chaotic bursting phenomena, where the process outputs exhibit intermittent short periods of bursts, followed by periods of quiescence (Anderson, 1985; Ydstie, 1997). Of course, if prior knowledge is available about the process to be identified, then PE in its standard form as an invertibility requirement on the information matrix may not be necessary, because the least-squares minimization is constrained to produce parameter values within a certain set (Lozano and Zhao, 1994).

Focusing on the essential role of PE in closed-loop identification, Genceli and Nikolaou (1996) introduced the simultaneous MPC and identification (MPCI) adaptive control paradigm. MPCI resorts to on-line optimization. An objective function over a moving horizon is minimized with respect to process inputs that satisfy all conventional MPC constraints and, in addition, a constraint that data be persistently exciting for the structure of the model being identified. Advantages of MPCI are that it does not require any explicit external dithering signals, PE is trivially dependent on the feasibility of the on-line optimization (by construction of the algorithm), constraints on process inputs are explicitly enforced, and constraints on process outputs are explicitly handled.

While the MPCI formulation may include a standard control objective, it provides wide flexibility for defining different objectives as well. For example, one could completely forego the on-line minimization of the sum of future square errors (which attempts to keep outputs at their setpoints) and simply attempt to keep outputs within specification limits during identification, while maximizing the information about the process to be identified. This brings us to the contribution of this work, namely the formulation of an MPCI variant with the following modification: process outputs are free to move away from setpoints, as long as they remain within specification bounds. The on-line objective is not to minimize the sum of square errors, but to maximize on-line the "magnitudes" (sum of the lower bounds on the minimum eigenvalues) of the information matrices over finite horizons. In that way, inputs to the process are allowed to excite the process as much as possible, for the generation of maximum parameter information, while the process output violates specification bounds as little as possible. Advantages of the proposed variant include the following:

(a) Easier controller design, through simplification of the on-line objective.

(b) Reduced computational load for the on-line optimization, by rendering the objective linear.

(c) Better satisfaction of realistic closed-loop identification goals, through replacement of setpoint tracking by output constraints.

The rest of this article is structured as follows: we first discuss the role of PE in identification and show how it naturally leads to MPCI. Next, we give a brief overview of the standard MPCI formulation. We then introduce a new MPCI variant and outline the solution to the corresponding on-line optimization problem via successive semidefinite programming (SSDP). We illustrate the new variant with computer simulations in which we make comparisons among a number of adaptive MPC variants and MPCI. Finally, we discuss further possibilities for MPCI.

2. PE and MPCI

The role of PE in identification has long been emphasized by a number of authors. We will discuss below a classic example found in several identification and adaptive control textbooks (Åström and Wittenmark, 1989, p.82; Ljung, 1987, p.365; Söderström and Stoica, 1989, p.26).

Example 1. The process to be identified is described by

$$y(k) = ay(k-1) + bu(k-1) + e(k) \tag{1}$$

where the disturbances $e(k)$ are independent and identically distributed (i.i.d.) random variables with zero mean (white noise). The parameters a and b have to be estimated while the process is controlled by a proportional regulator with gain K:

$$u(k) = -Ky(k) \tag{2}$$

A standard analysis of this example proceeds as follows: from the above two equations we get that the closed-loop behavior of the above system is

$$y(k) = (a - bK)y(k-1) + e(k) \tag{3}$$

which implies that all models $(\hat{a}, \hat{b})$ with

$$\hat{a} = a + \gamma K, \qquad \hat{b} = b + \gamma \tag{4}$$

where γ is an arbitrary scalar, will give the same input-output description of the process as the model (a, b) under the proportional controller feedback. Therefore, the parameters a and b cannot be estimated, although the process input u is persistently exciting, as a result of white noise going through a first-order process:

$$u(k) = (a - bK)u(k-1) - Ke(k) \tag{5}$$

Hence, *persistence of excitation is not a sufficient condition on the input in closed-loop experiments* (Ljung, 1987, p.366). The preceding statement, which is entirely correct, must be interpreted very carefully, because it might leave the impression that PE of process inputs is the only kind of excitation that would make identification possible. However, it is the richness (PE) of the data that appear in the information matrix that matters more. In fact, PE of process inputs is sufficient for process identification, whether in open or closed-loop, provided the process is modeled by an FIR model and noise is white. Let us explain the situation by looking at the above example in a slightly different way.

If least-squares identification of the above process were attempted, then it would require the minimization of

$$E\left[\sum_{i=1}^{k}\left(\hat{y}(i|i-1)-y(i)\right)^2\right]=E\left[\sum_{i=1}^{k}\left(\hat{a}y(i-1)+\hat{b}u(i-1)+e(i)-y(i)\right)^2\right] \quad (6)$$

with respect to the parameters $(\hat{a},\hat{b})$, at each time instant k. This minimization would, in turn, necessitate the solution to the linear system of equations

$$(\boldsymbol{X}^T\boldsymbol{X})\begin{bmatrix}\hat{a}\\ \hat{b}\end{bmatrix}=\boldsymbol{X}^T\begin{bmatrix}y(1)\\ \vdots\\ y(k)\end{bmatrix} \quad (7)$$

where

$$\boldsymbol{X}=\begin{bmatrix}y(0) & u(0)\\ \vdots & \vdots\\ y(k-1) & u(k-1)\end{bmatrix} \quad (8)$$

But the matrix $\boldsymbol{X}$, and consequently the matrix $\boldsymbol{X}^T\boldsymbol{X}$, has rank 1, because the second column of $\boldsymbol{X}$ is equal to the first column times $(-K)$, due to the presence of the proportional feedback controller, eqn. (2). Therefore, eqn. (7) does not have a unique solution but a one-dimensional family of solutions, captured by eqn. (4), and consequently identification (a,b) of is not possible.

It is evident that for the DARX model structure considered for this process, PE *of process inputs alone* is not immediately relevant to identification of the process. Instead, requirements should be placed on *both process inputs and outputs*, so that

(a) the pseudo-inverse of $\boldsymbol{X}$ can exist,

(b) the information matrix $\boldsymbol{X}^T\boldsymbol{X}$ is well conditioned,

(c) the matrix $\boldsymbol{X}$ and the noise vector $\boldsymbol{e}\hat{=}[e(0)\dots e(k-1)]^T$ are uncorrelated, i.e. $\lim_{k\to\infty}(1/k)\boldsymbol{X}^T\boldsymbol{e}=\boldsymbol{0}$, for unbiased estimates (Söderström and Stoica, 1989, p.186; Ljung, 1987, p.178).

The above requirements (a) and (b) correspond to the *strong* PE condition, referring to data of finite length (Goodwin and Sin, 1984, p.73), as contrasted to the *weak* PE condition that refers to data of infinite length. MPCI, as explained in the next section, explicitly addresses the above requirements (a) and (b), by forcing at each time step the information matrix $\boldsymbol{X}^T\boldsymbol{X}$ to be well conditioned, namely

$$\boldsymbol{X}^T\boldsymbol{X} \succeq \rho\boldsymbol{I} \succ \boldsymbol{0} \tag{9}$$

where the inequality symbol in eqn. (9) means that the information matrix $\boldsymbol{X}^T\boldsymbol{X}$ is positive semidefinite. In the sequel, the above inequality symbol will denote matrix definiteness whenever the inequality involves matrices on both sides. Note that maximization of ρ in eqn. (9) would also make small the parameter estimate bias $(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{e}$ (Ljung, 1987, eqn. (7.38)).

Continuing the analysis of the above example, let us assume that instead of the above DARX model, the following FIR model is used to approximately model the process:

$$y(k) = \sum_{i=1}^{m} b_i u(k-i) + e(k) \tag{10}$$

where e is erroneously assumed to be white noise. Note that the above use of an FIR model is applicable only to stable processes. Note, also, that the disturbance model does not agree with the disturbance model of eqn. (1). The process is again under the proportional feedback control of eqn. (2). Identification of the parameters $\{b_i\}_{i=1}^{m}$ according to the least-squares method would require minimization of

$$E\left[\sum_{i=1}^{k}\Big(\hat{y}(i|i-1) - y(i)\Big)^2\right] = E\left[\sum_{i=0}^{k}\Big(\sum_{j=1}^{m}\hat{b}_j u(i-j) + e(i) - y(i)\Big)^2\right] \tag{11}$$

with respect to the parameters $\{\hat{b}_i\}_{i=1}^{m}$. This minimization would, in turn, necessitate the computation of the pseudo-inverse $(\boldsymbol{U}^T\boldsymbol{U})^{-1}\boldsymbol{U}^T$ of the matrix

$$\boldsymbol{U} = \begin{bmatrix} u(-1) & \cdots & u(-m) \\ \vdots & \ddots & \vdots \\ u(k-1) & \cdots & u(k-m) \end{bmatrix} \tag{12}$$

The pseudo-inverse exists if $k \le m$ and the matrix $\boldsymbol{U}$ has full rank. The matrix $\boldsymbol{U}$, in turn, has full rank if the input u is persistently exciting, regardless of whether the identification experiment is conducted under open or closed-loop conditions and regardless of what kind of controller is used. (Closed-loop conditions will have an effect on the estimated parameter bias, since there will be correlation between noise and process input.) For the above closed-loop system (eqns. (1) and (2)), the input u is persistently exciting of any order, according to eqn. (5). Consequently, unique estimates of the Markov parameters $\{b_i\}_{i=1}^{m}$ can be obtained, even in the absence of signals externally introduced to the loop. Of course, according to standard asymptotic analysis, those estimates will be biased as

$$\hat{b}_1 = b - \frac{a}{K}, \quad \hat{b}_j = 0, \quad j = 2, 3, \ldots \tag{13}$$

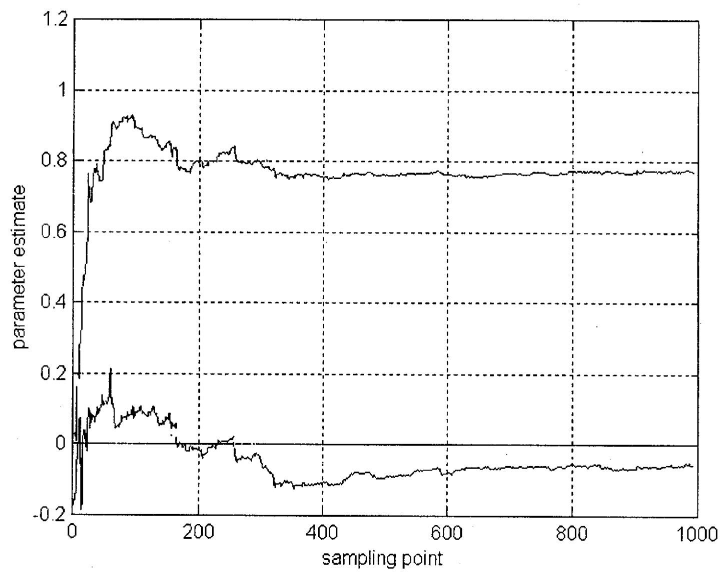

Fig. 1. Parameter convergence for eqn. (10), $a = 0.2$, $b = 1$, $K = 1$, $m = 3$.

in contrast to the ideal values of the Markov parameters $\{b_i\}_{i=1}^{m}$, which would be approximately $\{a^{i-1}b\}_{i=1}^{m}$ for large values of m. The bias can be reduced if external dithering signals are used. Figure 1, showing convergence of $\{\hat{b}_i\}_{i=1}^{m}$ without external dithering, verifies eqns. (13). ♦

As a final illustration of the above ideas, consider the following example:

Example 2. Consider least-squares identification of the process

$$y(k) = ay(k-1) + bu(k-1) + e(k) - ae(k-1) \tag{14}$$

controlled by proportional feedback as in eqn. (2). The process is parametrized by an FIR model as

$$y(k) = \sum_{i=1}^{m} b_i u(k-i) + e(k) \tag{15}$$

where the disturbance model of eqn. (15) is now correct. Then, the process input u is persistently exciting of any order (Ljung, 1987, p.362). Indeed, eqn. (16) multiplied by $-K$ yields

$$y(k) = -\sum_{i=1}^{m} b_i K u(k-i) - Ke(k) \Rightarrow u(z) = \frac{K}{1+\sum_{i=1}^{m} b_i K z^{-i}} e(z) \tag{16}$$

This means that, for large values of k, the matrix $\boldsymbol{U}$ in eqn. (12) will be invertible, hence the process will be identifiable. This is verified in the simulation of Fig. 2.

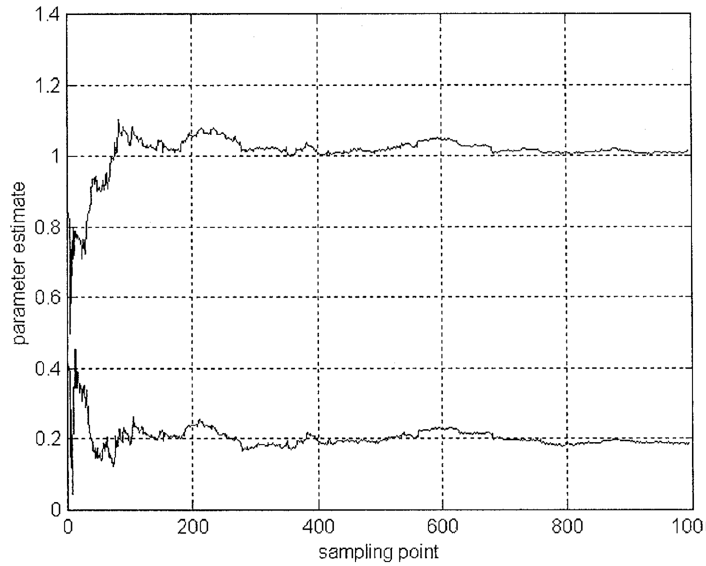

Fig. 2. Parameter convergence for eqn. (15), $a = 0.2$, $b = 1$, $K = 1$, $m = 2$.

In fact, this example demonstrates that external inputs are not necessary in closed-loop identification, although they may be extremely useful, e.g., in increasing the convergence rate. What is important is the quality of the data, for example the condition number of the information matrix, whichever way that result is achieved.

♦

Examples 1 and 2 above set the stage for the introduction of MPCI as an algorithm that explicitly requires the invertibility of the information matrix. Section 3 below gives a brief overview of MPCI, while Section 4 introduces a new MPCI variant.

3. MPCI On-Line Optimization Problem

Let the following linear time-varying model be used to represent the process at time k:

$$y(l|k) = \sum_{i=1}^{n} b_i(l|k)u(l-i|k) + d(l|k) \tag{17}$$

where $y(l|k)$ denotes the current or past output, $y(l)$, measured at time $l \leq k$, and the future output at time l predicted at time $k < l$, $u(l|k)$ stands for the implemented process input, $u(l)$, at time $l < k$, and the current or future potential input at time $l \geq k$, $d(l|k)$ signifies the estimate made at time k of the disturbance introduced at time l, and $b(l|k)$ denotes the estimate made at time k of a model coefficient at time l.

The disturbance d in eqn. (17) is the sum of a constant disturbance plus white noise. Based on the above, the MPCI on-line optimization can be formulated as follows (Genceli and Nikolaou, 1996):

MPCI Original On-Line Optimization Problem

Minimize

$$\sum_{i=1}^{M}\left[w_i\left(y(k+i|k)-y^{sp}\right)^2+r_i\Delta u(k+i-1|k)^2+v\varepsilon_i^2+h\mu_i^2\right] \tag{18}$$

with respect to $u(k|k),\ldots,u(k+M-1|k),\varepsilon_1,\ldots,\varepsilon_M$, subject to

$$u_{\max}\geq u(k+i-1|k)\geq u_{\min},\qquad i=1,2,\ldots,M \tag{19}$$

$$\Delta u_{\max}\geq \Delta u(k+i-1|k)\geq \Delta u_{\min},\qquad i=1,2,\ldots,M \tag{20}$$

$$y_{\max}+\mu_i\geq y(k+i|k)\geq y_{\min}-\mu_i,\quad \mu_i\geq 0,\quad i=1,\ldots,P \tag{21}$$

$$u(k+M+i|k)=u(k+i|k),\qquad i=0,\ldots,M-1 \tag{22}$$

$$y(k+i|k)=\phi(k+i|k)^T\overline{\theta}(k|k),\qquad i=1,2,\ldots,M \tag{23}$$

$$\overline{\theta}(k|k) = \left[\sum_{j=0}^{s-1}\lambda^j\phi(k-j|k)\phi(k-j|k)^T\right]^{-1} \times\left[\phi(k|k)\quad \phi(k-1|k)\quad\cdots\quad\phi(k-s+1|k)\right]\boldsymbol{Y}(k) \tag{24}$$

$$\sum_{j=0}^{s-1}\lambda^j\phi(k-j+i|k)\phi(k-j+i|k)^T\succeq(\rho_i-\varepsilon_i)\boldsymbol{I}\succ\boldsymbol{0},\quad i=1,2,\ldots,M \tag{25}$$

where

$$\Delta u(k+i-1|k)\hat{=}u(k+i-1|k)-u(k+i-2|k) \tag{26}$$

$$\boldsymbol{Y}(k)=\left[y(k)\quad y(k-1)\quad\cdots\quad y(k-s+1)\right]^T \tag{27}$$

$$\phi(k-j+i|k)=\left[u(k-j+i-1|k)\quad u(k-j+i-2|k)\quad\cdots\quad u(k-j+i-n|k)\quad 1\right]^T \tag{28}$$

v and h are weights on the softening variables ε_i and μ_i for the PE lower bounds ρ_i and output bounds $y_{\max}$ and $y_{\min}$, respectively, used only when the on-line optimization with $v=0$ and $\mu_i=0$ is infeasible; w is the weight on the square error between setpoint and predicted output; r is the weight on the move supression term; λ is a forgetting factor; and s refers to the length of the data set used for parameter estimation. In a typical MPC fashion (Prett and Garcia, 1988), the above optimization problem is solved at time k, and the optimal $u(k)$ is applied to the process.

Remarks:

- The main difference between MPC and MPCI is eqn. (25), the PE constraints. Switching between MPC and MPCI is simply a matter of including or excluding eqn. (25) in on-line optimization. Switching from MPC to MPCI is an interesting and formidable problem on its own right; to address that problem, additional information about the operating mode of a process may be helpful. For example, if a nonlinear process is moved from one steady-state operating condition to another, then a new local linear model may have to be identified. Information from statistics-based controller performance monitoring methods is also valuable, when such methods attribute poor controller performance to an inaccurate process model.
- The last entry for the vector ϕ in eqn. (28) is 1. This is because the nonzero average of the disturbance d is estimated as a component of the vector $\overline{\theta}$ in eqn. (24).
- It is evident from eqns. (24) and (25) that an "average" parameter estimate $\overline{\theta}$ is used for output prediction.
- Equation (22) implies that MPCI has an M-periodic sequence of future inputs $u(k+i|k)$.
- The feasibility of the on-line optimization problem is guaranteed by the existence of the softening variables ε_i and μ_i. To reduce the dimensionality of the on-line optimization problem, one may assume that $\mu_i = \mu$ and $\varepsilon_i = \varepsilon$. This will result in minor loss of controller flexibility, because output constraints are important only for the first few steps of the prediction horizon. It must be noted that the softening variables ε_i and μ_i are used when the on-line optimization problem with hard output constraints is infeasible. The weights v and h on the variables ε_i and μ_i should be selected a few orders of magnitude above the weight on the output error, to ensure that softening will violate constraints (21) and (25) as little as possible.
- The parameter s can be selected freely, as long as it satisfies the inequality

$$s > n \tag{29}$$

- The choice of the weights in the objective function of eqn. (18) is important for a balance between control and identification. While sensible choices can be made on the basis of engineering arguments, the issue will be simplified with the introduction of a new MPCI variant in Section 4.
- Genceli and Nikolaou (1996) developed a successive semi-definite programming (SSDP) algorithm with guaranteed convergence to a local optimum of the on-line optimization problem (Shouche *et al.*, 1996). The computational complexity of SSDP is not much higher than the complexity of quadratic programming (QP) used in standard MPC. SDP and SSDP are briefly explained in Section 5.
- Notice that only $u(k|k)$ among all decision variables $\{u(k|k), \ldots, u(k|k+M-1)\}$ appears in inequality (25), as shown in Fig. 3. The important implication of this observation is that PE and maximal information matrix are trivially guaranteed

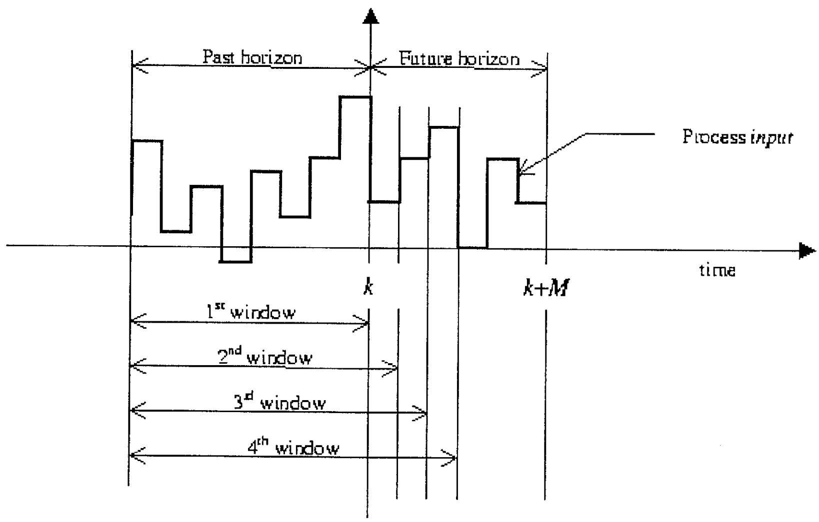

Fig. 3. MPCI moving horizon.

for the closed loop, regardless of the behavior of the true process, as long as the on-line optimization problem has a feasible solution with $\rho_1 > 0$. Consequently, parameter convergence can also be easily established.

4. Replacing Output Regulation by Output Constraints

Since MPCI relies on on-line optimization, it provides wide flexibility for formulating various constraints and defining objectives other than the standard minimization of 2-norms. For example, one could completely forego the minimization of the sum of square errors (which attempts to keep outputs at their setpoints) and simply attempt to keep outputs within specification bounds during identification, while maximizing the lower eigenvalue of the corresponding information matrices. The control objective may be loosened or tightened by adjusting the output constraint bounds. In addition, the MPCI formulation trivially guarantees, by construction, that data are going to be as informative as possible. This flexibility allows us to develop a new mathematical formulation of realistic engineering goals associated with the identification of a process controlled by constrained MPC. To the best of our knowledge, all past dual control formulations have included a regulation related term (sum of square errors), which limits the flexibility of dual control. Note that the PE constraints, eqn. (25), are not convex. These are the only MPCI constraints that are not convex, because of the nature of MPCI as a PE enforcing algorithm. All other constraints as well as the objective function may capture a very wide class of engineering requirements without sacrificing convexity (Boyd *et al.*, 1994). Below we are showing two MPCI variants that capitalize on the MPCI flexibility discussed above, by replacing output setpoint-tracking with output constraints.

4.1. MPCI Variant P1

As discussed above, in this MPCI variant the process output is free to move away from its setpoint but is required to remain within specification bounds. The on-line optimization problem is not to minimize a quadratic objective function involving inputs and states, but to maximize the sum of lower bounds on the minimum eigenvalues of the information matrices over a finite horizon. In that way, inputs to the process are allowed to excite the process highly enough to generate as much information on process parameters as possible, while the process is outside specifications as little as possible. The on-line optimization problem becomes as follows:

$$\text{minimize} \quad \sum_{i=1}^{M} \rho_i - h \sum_{i=1}^{P} \mu_i \tag{30}$$

with respect to $u(k|k), \ldots, u(k+M-1|k)$, $\rho_1, \ldots, \rho_M$, $\mu_1, \ldots, \mu_P$, subject to the constraints of eqns. (19) to (24) and the unsoftened PE constraints

$$\sum_{j=0}^{s-1} \lambda^j \phi(k-j+i|k)\phi(k-j+i|k)^T \succeq \rho_i \boldsymbol{I} \succ \mathbf{0}, \quad i = 1, \ldots, M \tag{31}$$

For MPCI Variant P1 we select

$$s = M \tag{32}$$

Remarks:

- As seen from eqns. (30) and (31), the above MPCI variant maximizes the lower bounds on the minimum eigenvalues of information matrices. In this way, inputs are forced to satisfy a PE constraint, in order to generate as much parameter information as possible.
- The choice $s = M$, eqn. (32), corresponds to a forgetting factor that is identically zero for the data older than M time steps.
- The engineering objective of keeping the output y closed to its setpoint is expressed by the output constraint in eqn. (21).
- Since a poor model is used to predict future values of the output y, at least at the initial stages of identification, care must be taken to express the output constraints conservatively enough, to compensate for inaccuracies in the prediction of y. A simple approach to this issue is to tighten the upper and lower bounds on y, $y_{\max}$ and $y_{\min}$, respectively. A more sophisticated approach, relying on chance-constraint formulation of output constraints is presented in (Schwarm *et al.*, 1998).
- In lieu of eqn. (24), one can easily use the recursive least squares (RLS) algorithm. $\boldsymbol{P}(0)$, the covariance of $\hat{\theta}(0)$, is usually chosen sufficiently large at the beginning of

identification. Parameter convergence using the recursive least squares algorithm is guaranteed if the following condition is satisfied:

$$\lim_{k\to\infty} \lambda_{\min}\left(\sum_{j=1}^{k-1}\phi(j)\phi(j)^T\right) = \infty \tag{33}$$

4.2. MPCI Variant P2

The difference between this MPCI variant and P1 is that instead of requiring PE in a moving horizon with fixed window length, as in eqn. (31), we enforce PE over a series of windows of increasing length, as follows:

$$\sum_{j=0}^{k-1+i} \lambda^j \phi(k-j+i|k)\phi(k-j+i|k)^T \succeq \rho_i \boldsymbol{I}, \quad i = 1,\ldots,M \tag{34}$$

The common starting point of the PE horizons (Fig. 3) is at the beginning of identification.

Remarks:

- Notice that the upper limit of the summation in eqn. (34) is $k-1+i$. This increases the identification horizon length at each time step k, as is standard in recursive least squares (RLS).
- Since this variant retains the old data and keeps adding more new data, it has to be used in the cases where the parameters do not change during identification. In that case there is valuable information in all past data. In contrast, Variant P1 must be used in the cases where parameters change with time, and distant past information has to be discarded.
- If eqn. (24) is replaced by recursive least squares (RLS), then it is straightforward to show that the PE constraints in Variant P2, eqn. (34), purports to generate data that do not result in singularities in RLS.
- Both sides of eqn. (34) can be multiplied by $1/k$, to prevent excessively large numbers for large k.

5. Numerical Solution of the MPCI On-Line Optimization Problem

The inclusion of PE constraints in the MPCI on-line optimization generates a non-convex problem, because the PE constraints are non-convex quadratic matrix inequalities (QMI). While the global optimum of this problem may be hard to find, a local optimum can be easily found. Genceli and Nikolaou (1996) developed an algorithm that successively employs semidefinite programming (SDP), to converge to a local optimum. Shouche *et al.* (1998) showed that convergence of the successive SDP (SSDP) of Genceli and Nikolaou (1996) is guaranteed. We give a brief overview of that algorithm below, and illustrate how it can be applied to the MPCI variants of this work.

5.1. Approximation of QMI by LMI

To circumvent the nonconvexity problem resulting from the PE constraints, we employ the following inequality (Genceli and Nikolaou, 1996):

$$\sum_{j=0}^{r} \phi^j(k-j+i|k)\phi(k-j+i|k)^T \succeq L_i(\boldsymbol{U}(k)), \quad i=1,\ldots,M \tag{35}$$

where

$$\begin{aligned} L_i(\boldsymbol{U}(k)) \;\hat{=}\; & \sum_{j=0}^{r} \lambda^j \phi^*(k-j+i|k)\phi(k-j+i|k)^T \\ & + \sum_{j=0}^{r} \lambda^j \phi(k-j+i|k)\phi^*(k-j+i|k)^T \\ & + \sum_{j=0}^{r} \lambda^j \phi^*(k-j+i|k)\phi^*(k-j+i|k)^T \end{aligned} \tag{36}$$

$$\boldsymbol{U}(k) = \begin{bmatrix} u(k|k) & u(k+1|k) & \cdots & u(k+M-1|k) \end{bmatrix}^T \tag{37}$$

and the upper limit, r, of the summation in eqn. (35) is equal to $M-1$ in eqn. (31), and $k+i-1$ in eqn. (34), for the MPCI Variants P1, and P2, respectively. The vectors $\phi^* \in \mathbb{R}^{n+1}$ in eqn. (36) can be thought of as points of linearization (Genceli and Nikolaou, 1996).

Therefore, by virtue of eqns. (35), the QMI in eqns. (31) or (34) are satisfied if the inequalities

$$L_i(\boldsymbol{U}(k)) \succeq \rho_i \boldsymbol{I}, \quad i=1,2,\ldots,M \tag{38}$$

are satisfied. The inequalities in eqn. (38) can be easily shown to be linear matrix inequalities (LMI), i.e., they have the general form

$$\boldsymbol{G}(\boldsymbol{z}) \;\hat{=}\; \boldsymbol{G}_0 + \sum_{i=1}^{q} z_i \boldsymbol{G}_i \succeq \boldsymbol{0} \tag{39}$$

where $\boldsymbol{z} \in \mathbb{R}^q$ is a variable, and the symmetric matrices $\boldsymbol{G}_i = \boldsymbol{G}_i^T \in \mathbb{R}^{N\times N}$, $i = 0,\ldots,q$, are given. It can be further shown (Genceli and Nikolaou, 1996) that the M LMI's in eqn. (38) can be written as a single LMI of the form

$$\underbrace{\begin{bmatrix} \boldsymbol{W}_{1,0} & \boldsymbol{0} & \boldsymbol{0} \\ 0 & \ddots & \boldsymbol{0} \\ 0 & 0 & \boldsymbol{W}_{M,0} \end{bmatrix}}_{\boldsymbol{F}_0} + \sum_{t=0}^{M-1} u(k+t) \underbrace{\begin{bmatrix} \boldsymbol{C}_{1,t} & \boldsymbol{0} & \boldsymbol{0} \\ 0 & \ddots & \boldsymbol{0} \\ 0 & 0 & \boldsymbol{C}_{M,t} \end{bmatrix}}_{\boldsymbol{F}_{t+1}} + \sum_{i=1}^{M} \rho_i \underbrace{\begin{bmatrix} \boldsymbol{T}_{1,i} & \boldsymbol{0} & \boldsymbol{0} \\ 0 & \ddots & \boldsymbol{0} \\ 0 & 0 & \boldsymbol{T}_{M,i} \end{bmatrix}}_{\boldsymbol{F}_{M+i}} \succ \boldsymbol{0} \tag{40}$$

Details can be found in (Genceli and Nikolaou, 1996).

Therefore, one can find a sub-optimal solution to Problems P1 or P2, if one solves the optimization problem of the following form:

$$\text{minimize} \quad \left(-\sum_{i=1}^{M} \rho_i + h \sum_{i=1}^{P} \mu_i \right) \tag{41}$$

with respect to $\boldsymbol{x}$, subject to

$$\boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b} \tag{42}$$

$$\boldsymbol{F}_0 + \sum_{i=1}^{2M+P} x_i \boldsymbol{F}_i \succeq \boldsymbol{0} \tag{43}$$

where

$$\boldsymbol{x} = \begin{bmatrix} u(k|k) & u(k+1|k) & \cdots & u(k+M-1|k) & \rho_1 & \cdots & \rho_M & \mu_1 & \cdots & \mu_P \end{bmatrix}^T \tag{44}$$

and the inequality of eqn. (42) captures all MPCI inequality constraints excluding the PE constraints. The above optimization problem in eqns. (41) to (43) can finally be reformulated as the following SDP problem:

MPCI On-line SDP Problem

Minimize

$$\boldsymbol{c}^T \boldsymbol{x} \tag{45}$$

with respect to $\boldsymbol{x}$, subject to

$$\underbrace{\begin{bmatrix} -\mathrm{Diag}\,(\boldsymbol{b}) & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{F}_0 \end{bmatrix}}_{\boldsymbol{G}_0} + \sum_{i=1}^{2M+P} x_i \underbrace{\begin{bmatrix} \mathrm{Diag}\,(\boldsymbol{a}_i) & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{F}_i \end{bmatrix}}_{\boldsymbol{G}_i} \succeq \boldsymbol{0} \tag{46}$$

where $\boldsymbol{A} = [\boldsymbol{a}_1 \ \boldsymbol{a}_2 \ \cdots \ \boldsymbol{a}_{2M+P}]$, $\boldsymbol{a}_i$, is a column vector; $\mathrm{Diag}\,(\boldsymbol{v})$ is the diagonal matrix with diagonal elements equal to the entries of the vector $\boldsymbol{v}$; $\boldsymbol{c} = \begin{bmatrix} 0 & \cdots & \underset{M\text{-th}}{0} & -1 & \cdots & \underset{2M\text{-th}}{-1} & 1 & \cdots & \underset{(2M+P)\text{-th}}{1} \end{bmatrix}^T \in \mathbb{R}^{2M+P}$; $\boldsymbol{x} \in \mathbb{R}^{2M+P}$; $\boldsymbol{F}_i$, $i = 0, \ldots, 2M$, are defined in eqn. (40); $\boldsymbol{F}_i = \boldsymbol{0}$, $i = 2M+1, \ldots, 2M+P$.

5.2. Semidefinite Programming (SDP)

Optimization problems involving minimization of a linear function subject to an LMI are called *Semidefinite Programming (SDP)* problems (Vandenberghe and Boyd, 1996). These are convex optimization problems of the form:

$$\text{minimize} \quad \boldsymbol{c}^T \boldsymbol{z} \tag{47}$$

with respect to z, subject to

$$G_0 + \sum_{i=1}^{q} z_i G_i \succeq 0 \tag{48}$$

The solution to SDP problems can be obtained by powerful algorithms like interior point methods. In our simulations, solutions are obtained by a Merhotra-type primal-dual predictor-corrector interior-point algorithm for semidefinite programming (Alizadeh *et al.*, 1997).

It should be noted that, by its structure, the SDP problem corresponding to MPCI is guaranteed to have a feasible solution.

5.3. Successive SDP

To obtain an SDP problem of eqns. (45) and (46) at a given time step k, the PE constraints, eqns. (31) or (34), were substituted by LMI after linearization around the vectors $\phi^* \in \mathbb{R}^{n+1}$. If ϕ^* are poorly chosen, then the solution to the on-line SDP problem may be far from the original MPCI nonconvex problem (Variants P1 and P2). To improve the solution to the MPCI on-line problem, one can iteratively refine the vectors ϕ^* and successively solve each corresponding SDP problem. Refined ϕ^* contain values of u that are optimal for the SDP problem of the previous iteration corresponding to the previous vector ϕ^* (Genceli and Nikolaou, 1996). It should be noted that the solution to each successive SDP problem is a feasible solution to the original problem. We summarize the SSDP procedure below.

Successive SDP for the solution of the MPCI on-line problem

Step 1. Compute the matrices $\{G_i,\ i = 1, 2, \ldots, 2M + P\}$ in eqn. (46).

Step 2. Select
$\phi^*(k - j + i|k)$, $j = 0, \ldots,$ $\{M - 1$ for Variant P1; $M + i - 1$ for P2, $k + i - 1$ for P3$\}$ $i = 1, \ldots, M - 1$
A good choice is:

$$\phi^*(k - j + i|k) = \begin{bmatrix} \alpha_1 & \cdots & \alpha_n & 1 \end{bmatrix}^T \tag{49}$$

where

$$\alpha_l = \begin{cases} u(k - j + i - l) & \text{if } j - i + l \geq 1 \text{ (past } u\text{)} \\ u_{\mathrm{SDP}}(k - j + i - l|k - 1) & \text{if } j - i + l < 1 \text{ (SDP solution at } k-1\text{, but not implemented)} \end{cases} \tag{50}$$

Step 3. Solve the SDP problem (45), (46). Let the optimal solution be x

$$\left\{u_{sp}(k|k),\ldots,u_{sp}(k+M-1),\ \rho_1,\ldots,\rho_M,\ \mu_1,\ldots,\mu_P\right\} \tag{51}$$

Step 4. Update ϕ^* as follows:

$$\phi^*_{\text{new}}(k-j+i|k) = \begin{bmatrix}\delta_1 & \cdots & \delta_n & 1\end{bmatrix}^T \tag{52}$$

where

$$\delta_l = \begin{cases} u(k-j+i-l) & \text{if } j-i+l \geq 1 \text{ (past } u\text{)} \\ u_{\text{SDP}}(k-j+i-l|k) & \text{if } j-i+l < 1 \text{ (solution at } k\text{)} \end{cases} \tag{53}$$

Step 5. If $(c^T x_{\text{old}} - c^T x_{\text{new}}) \geq \kappa_1$ and $\|\phi^*_{\text{old}} - \phi^*_{\text{new}}\| \geq \kappa_2$ then $x_{\text{old}} = x_{\text{new}}$ and $\phi^*_{\text{old}} = \phi^*_{\text{new}}$, go to Step 3. Otherwise, go to Step 6.

Step 6. Implement the first input move $u(k|k)$ to the process.

Step 7. Identify the process parameter using eqn. (24) or RLS.

Step 8. Update the output prediction using eqn. (23).

Step 9. Let $k \leftarrow k+1$ and go to Step 1.

6. Illustrative Examples

Example 3. Let the real behavior of a linear process be described by the equation

$$y(k) = u(k-1) + 0.5u(k-2) + 0.2u(k-3) + 0.1u(k-4) + d(k) + w(k) \tag{54}$$

where d is a deterministic disturbance and w is white noise with zero mean and standard deviation equal to 0.01. The process input u must satisfy the constraints

$$-0.4 \leq u(k) \leq 0.1 \tag{55}$$

at all times k. Assume that the linear model

$$\begin{aligned} y(k+i|k) &= 1.1u(k+i-1|k) + 0.55u(k+i-2|k) \\ &\quad + 0.22u(k+i-3|k) + 0.11u(k+i-4|k) + d(k+i|k) \end{aligned} \tag{56}$$

is available for the above process from previous data.

The process experiences a step setpoint change

$$y^{sp} = -0.3 \tag{57}$$

at time $k = 0$. After this setpoint change, we want the output y to be constrained as

$$-0.4 \leq y(k) \leq -0.25 \tag{58}$$

to meet specifications. The process also experiences the following upsets:

Period 1. $(0 \le k \le 25)$ The system is upset by the step disturbance

$$d(k) = 0.1 \tag{59}$$

starting at time $k = 0$;

Period 2. $(26 \le k \le 100)$ At time $k = 26$ the real process changed as follows:

$$y(k) = -0.3u(k-1) - 0.2u(k-2) - 0.1u(k-3) + 0.5u(k-4) + d(k) + w(k) \tag{60}$$

The system is upset by the step disturbance

$$d(k) = -0.4 \tag{61}$$

We use the following four control schemes to control the above process:

A. MPC,

B. MPC with adaptation,

C. MPC with adaptation and external dithering, and

D. MPCI Variant P1,

The controller parameter values are shown in Table 1.

Table 1. Values of controller parameters for Example 3.

Parameter	Symbol	Value for MPC Case C	Value for MPCI Case D
Control horizon length	M	8	8
Optimization horizon length	P	12	12
Move suppression term weight	r_i	0.1	not applied
Output error term weight	$w_i,\ i = 1, \ldots, M+n$	1	not applied
Forgetting factor	λ	1	1
Weight of constraint softening variable, ε	h	not applied	100
Standard deviation of dithering signal		0.25	not applied

Case A: MPC

The conventional MPC resulted in saturation of the process input and never recovered as illustrated in Fig. 4.

Case B: MPC with adaptation

Adaptation was added to the MPC scheme. Adaptation began at $k = 5$ after sufficient data were collected. Process inputs reached saturation, thus resulting in a badly conditioned information matrix. Thus parameters could not be identified properly. This scheme did not recover, either (Figs. 5 and 6).

Case C: MPC with adaptation and dithering on the process input

The process input was dithered through addition of a normally distributed signal with zero mean and standard deviation of 0.25. The intention was that dithering would result in PE of process inputs. Although this scheme did better than MPC with adaptation, parameters did not converge to real process coefficients. In addition, the process output y did not stay within the specified bounds (eqn. (58)) most of the time (Figs. 7 and 8).

Case D: MPCI

After standard MPC was implemented for the first 25 time steps, MPCI was turned on at the 26-th time step, when the process changed. The process output y stayed within the specified bounds (eqn. (58)). Adaptation began at $k = 32$. MPCI correctly identified the new process parameters and disturbance after sufficient data were collected, i.e. when $k > 32$ (Figs. 9 and 10). ♦

Example 4. Consider again the real behavior of a linear process described by an equation similar to (54), with the noise variable $w(t)$ being colored as

$$w(k) = e(k) + 0.8e(k-1) \tag{62}$$

where e is white noise with standard deviation $\sigma_e = 0.01$ and $\sigma_e = 0.1$. Table 2 shows estimates at the final point of the simulations. As expected, estimates are somewhat biased. The effect of colored noise on MPCI and possible modifications need to be further investigated.

Table 2. Estimates of Example 4.

Parameter	Parameter estimate for $\sigma_e = 0.01$	Parameter estimate for $\sigma_e = 0.1$
$\hat{d}$ ($d = -0.4$)	-0.3992	-0.3929
$\hat{b}_1$ ($b_1 = -0.3$)	-0.2966	-0.1900
$\hat{b}_2$ ($b_1 = -0.2$)	-0.1950	-0.2319
$\hat{b}_3$ ($b_1 = -0.1$)	-0.1005	-0.1325
$\hat{b}_4$ ($b_1 = 0.05$)	0.0481	-0.0373

♦

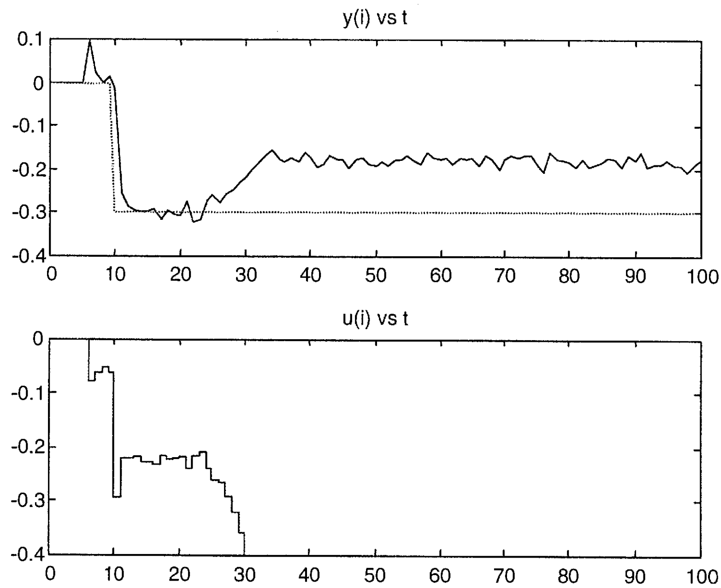

Fig. 4. Process output and input for closed loop with MPC, Example 3.

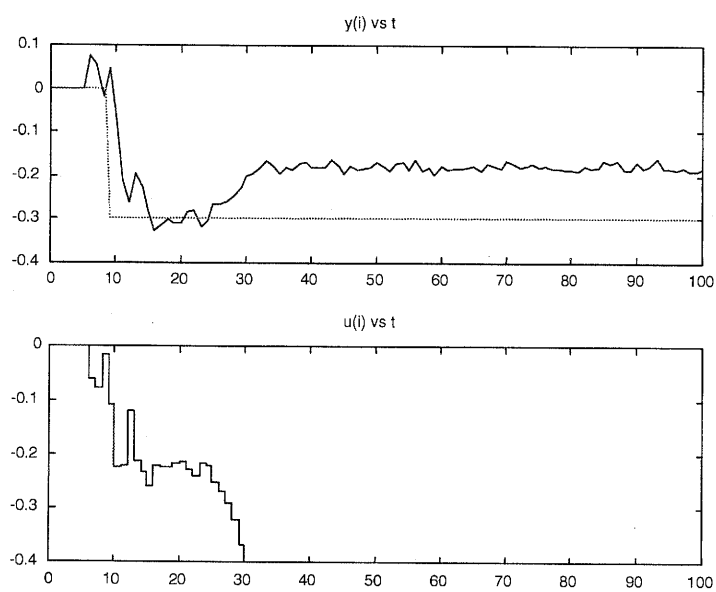

Fig. 5. Process output and input for closed loop with adaptive MPC, Example 3.

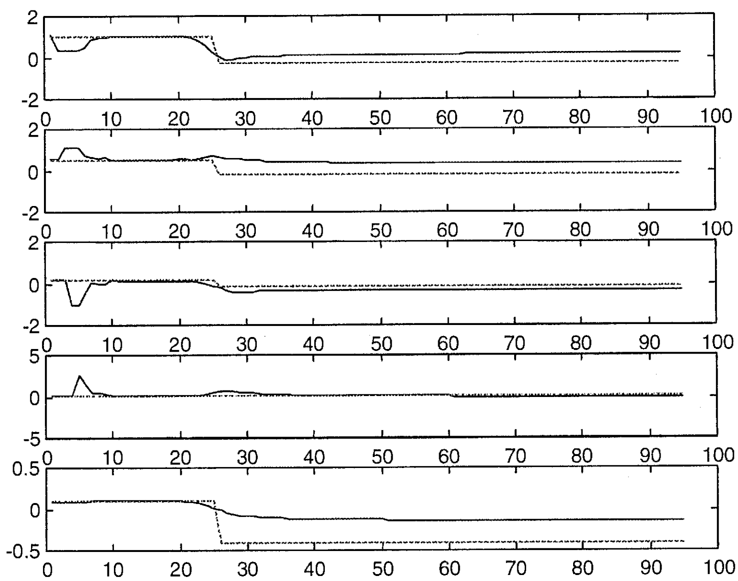

Fig. 6. Process model parameters and disturbance estimates for closed loop with adaptive MPC, Example 3 (dashed lines are real values, continuous line are estimates).

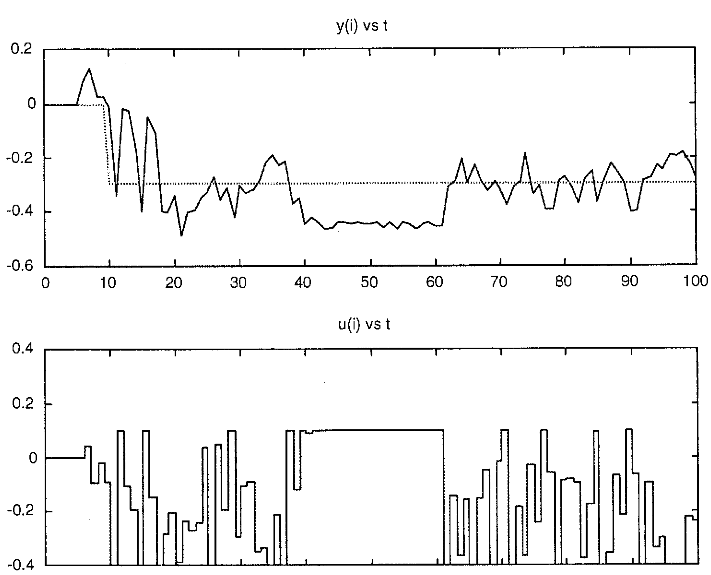

Fig. 7. Process output and input for closed loop with adaptive MPC and external dithering, Example 3.

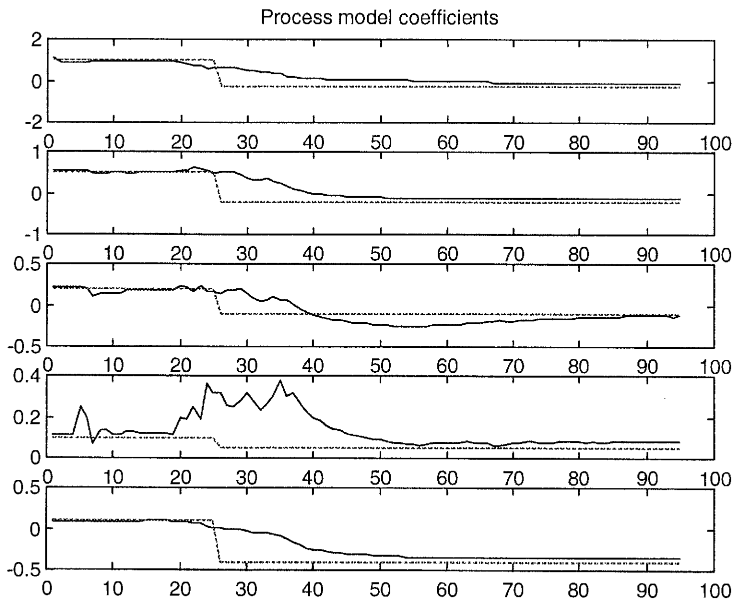

Fig. 8. Process model parameters and disturbance estimates for closed loop with adaptive MPC and dithering, Example 3 (dashed lines are real values, continuous line are estimates).

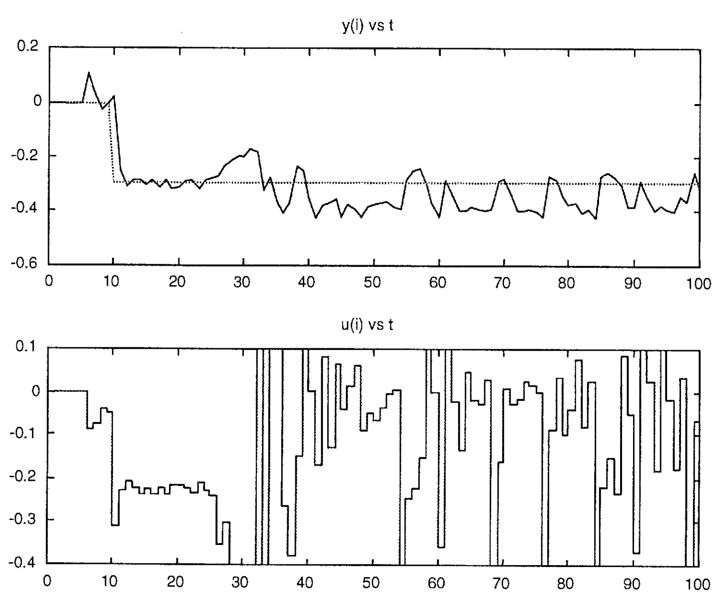

Fig. 9. Process output and input for closed loop with MPCI Variant P1, Example 3.

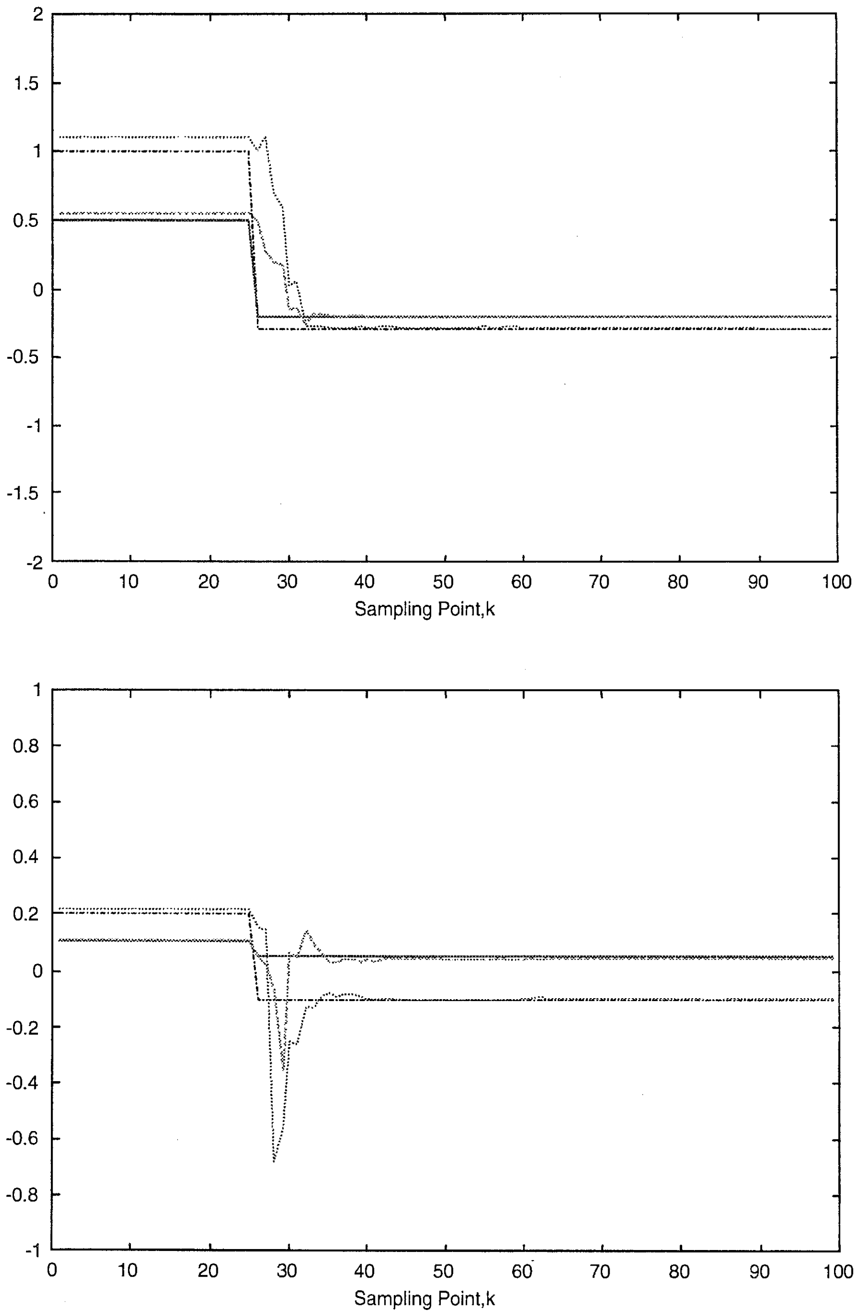

Fig. 10. Process model parameters and disturbance estimates for closed loop with MPCI, Example 3 (dashed lines are real values, continuous line are estimates).

Example 5. Consider the linear process

$$y(k) = u(k-1) + 0.2u(k-2) + 0.04u(k-3) + d(k) + w(k) \tag{63}$$

where d is a deterministic disturbance and w is white noise with zero mean. The process input must satisfy the constraints

$$-1 \leq u(k) \leq 1 \tag{64}$$

$$-0.5 \leq y(k) \leq 0.5 \tag{65}$$

at all times k. Note that the bounds on the output y are tighter than the actual specification bounds, to account for model uncertainty during identification. Assume that the linear model

$$y(k+i|k) = 2u(k+i-1|k) + 0.1u(k+i-2|k) + d(k+i|k) \tag{66}$$

is available for the above process from prior information. The process is disturbed by white noise w with zero mean and $\sigma_w = 0.1$, and by a step disturbance of magnitude 1.5, entering the system at time $k = 30$. MPCI (Variant P2) is compared to the following adaptive MPC scheme: Minimize

$$\sum_{i=0}^{P-1} \Big(u(k+i|k) - u_{sp}(k+i)\Big)^2 + h\sum_{i=1}^{P} \mu_i^2 \tag{67}$$

with respect to u, subject to

$$y_{\min} - \mu_i \leq y(k+i|k) \leq y_{\max} + \mu_i, \quad i = 1, \ldots, P \tag{68}$$

$$u_{\min} \leq y(k+i|k) \leq u_{\max}, \qquad i = 1, \ldots, P-1 \tag{69}$$

and eqns. (24) and (25) corresponding to MPCI Variant P2, where u_{sp} is a predetermined persistently exciting input sequence, and h is the weighting factor on the output constraint softening variable. The first term in eqn. (67) purports to force the process input u to be persistently exciting. The second term is used only when the on-line optimization is infeasible with $h = 0$. Note that the only differences between the MPCI and adaptive MPC schemes of this example are (a) the objective functions, eqn. (30) vs. eqn. (67), and (b) the presence of PE constraints in MPCI. Note, also, that both the schemes require the output y not to track a setpoint, but merely to stay within specification bounds. Values for controller parameters are given in Table 3.

Table 3. Values of controller parameters for Example 5.

Parameter	Symbol	Value for Adaptive MPC	Value for MPCI
Control horizon length	M	5	5
Optimization horizon length	P	5	5
Forgetting factor	λ	1	1
Weight of constraint softening variables, ε, μ	h	1000	100

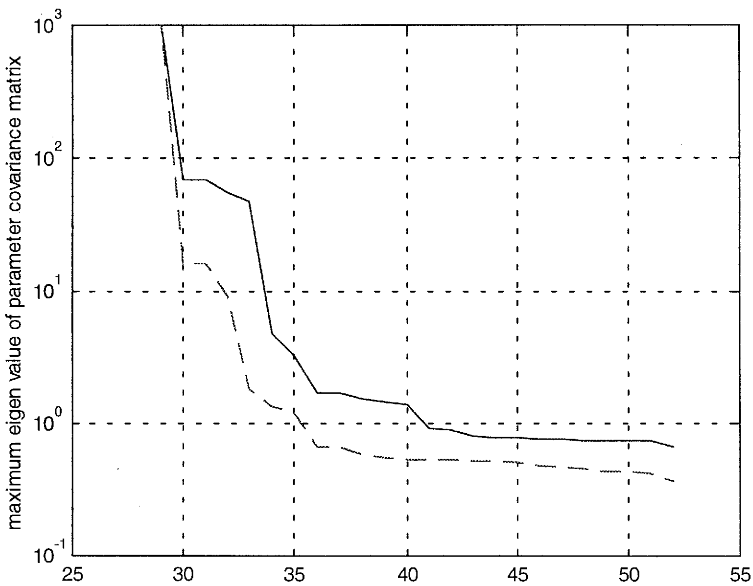

Fig. 11. Comparison of the size of parameter covariance estimates for adaptive MPC (solid line) and MPCI (dashed line) for Example 5.

For both the control schemes, the process is under MPC until $k = 25$, and under the above adaptive MPC until $k = 29$. The reason for not switching on MPCI immediately is discussed in (Genceli and Nikolaou, 1996). At time $k = 30$ a step disturbance of magnitude 0.2 enters the system.

Figure 11 compares estimates of the parameter covariance matrices corresponding to MPCI and MPC. The clear superiority of MPCI is evident. What is even more important is that although both the controllers force the process output beyond its bounds (as expected, because of uncertainty in future output predictions due to model inaccuracies) MPCI produces a *lower* sum of squares of output constraint violation errors than the specific adaptive MPC. The reason is that MPCI selects inputs that maximize information in the presence of an external step disturbance, while attempting to keep the process output within bounds. Adaptive MPC, on the other hand, attempts to reconcile disturbance rejection and identification in a way that is not optimal, because it forces the process input to follow a signal that would be persistently exciting, hence optimal, in the absence of disturbance. ◆

7. Conclusions

In this work we discussed the basic philosophy behind MPCI and proposed a new MPCI variant. In this variant, process outputs are free to move away from setpoints, as long as they remain within specification bounds. Process inputs, on the other hand, are constrained to excite the process as much as possible, for the generation of

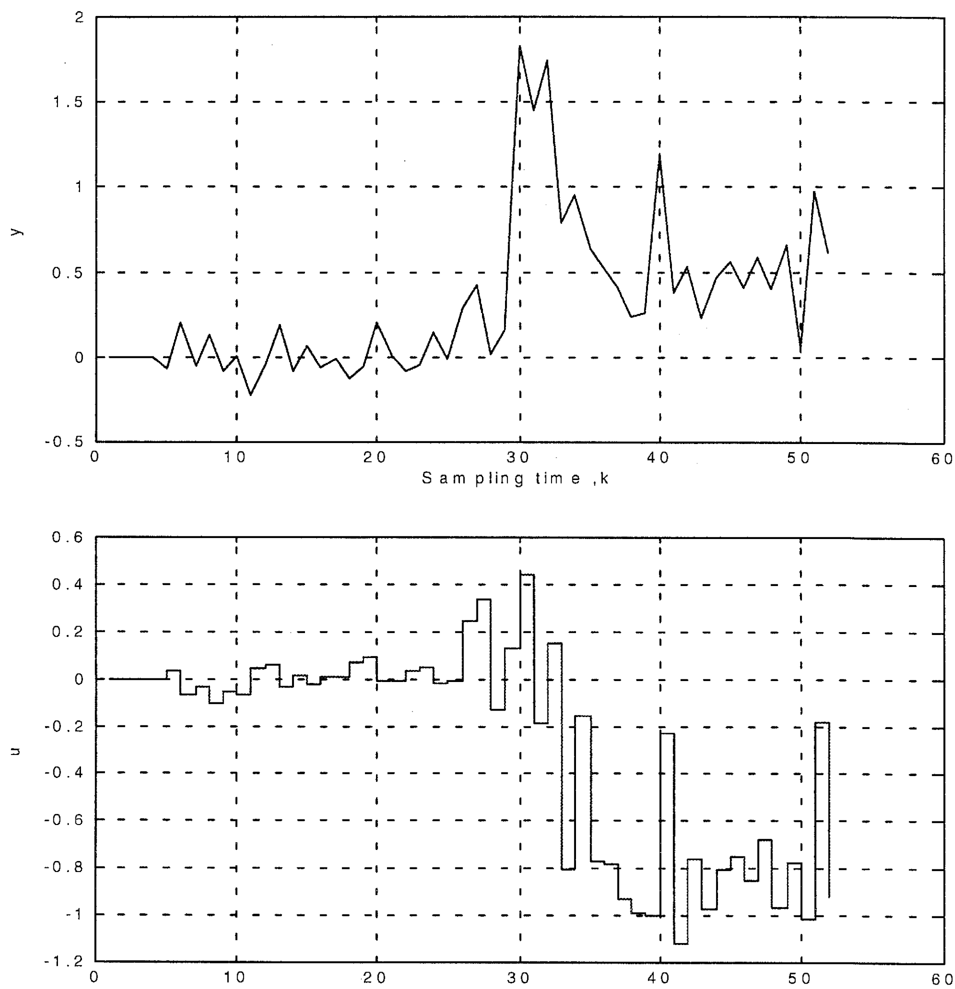

Fig. 12. Process output and input for the closed loop with adaptive MPC, Example 5.

maximum parameter information, while process outputs violate specification bounds as little as possible. By its construction, the proposed algorithm reduces the analysis of closed-loop PE enforcement to a feasibility problem. Advantages of the proposed variant include

- easier controller design, through simplification of the on-line objective,
- reduced computational load for the on-line optimization, by rendering the objective linear (Boyd *et al.*, 1994),
- better satisfaction of realistic closed-loop identification goals, through the replacement of setpoint tracking by output constraints.

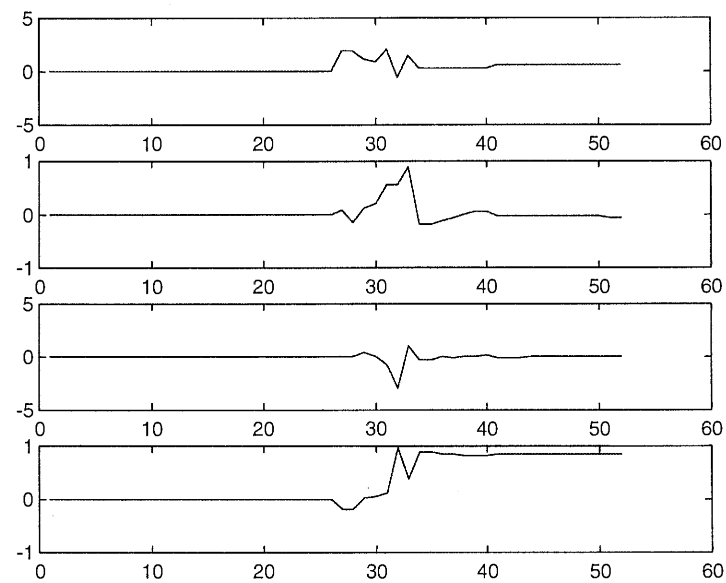

Fig. 13. Parameter convergence for closed-loop identification with adaptive MPC, Example 5.

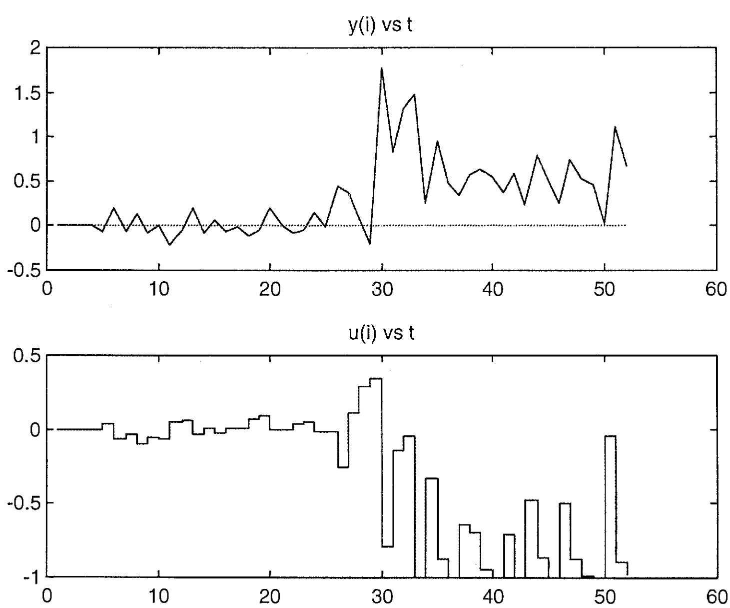

Fig. 14. Process output and input for the closed loop with MPCI, Example 5.

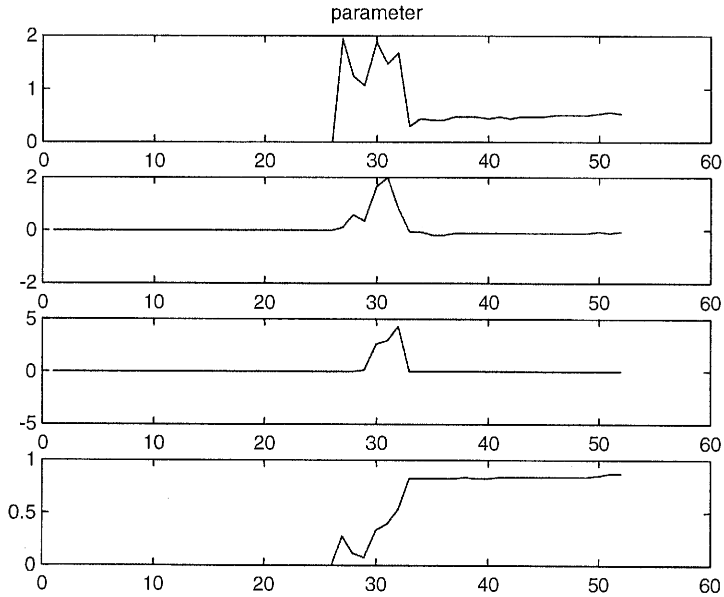

Fig. 15. Parameter convergence for closed-loop identification with MPCI, Example 5.

The merits of the proposed approach were illustrated through two simulation examples, in which the superiority of MPCI over various versions of adaptive MPC (essentially the only alternative for constrained control) was demonstrated.

While the basis philosophy of MPCI is simple, namely inclusion of PE constraints in on-line optimization, there are its many facets that need to be studied, along the lines of the rich literature on system identification.

References

Alizadeh F., Haeberly J.P., Nayakkankuppam M.V., Overton M.L. and Schmieta S. (1997): *SPD Pack User's Guide.* — NYU Computer Science Dept. Technical Report.

Anderson B.D.O. (1985): *Adaptive systems, lack of persistency of excitation and bursting phenomenon.* — Automatica, Vol.21, No.3, pp.247–258.

Anderson and Johnson (1982): *Exponential convergence of adaptive identification and control algorithms.* — Automatica, Vol.18, No.1, pp.1–13.

Åström K.J. and Wittenmark B. (1989): *Adaptive Control.* — Reading, MA: Addison Wesley.

Boyd S. and El Ghaoui L., Feron E. and Venkataramanan B. (1994): *Linear Matrix Inequalities in System and Control Theory.* — Philadelphis: Society for Industrial and Applied Mathematics.

Genceli H. and Nikolaou M. (1996): *A new approach to constrained predictive control with simultaneous model identification.* — AIChE J., Vol.42, No.10, pp.2857–2868.

Goodwin G.C. and Sin K.S. (1984): *Adaptive Filtering, Prediction and Control.* — Englewood Cliffs, N.J.: Prentice-Hall.

Lozano R. and Zhao X.H. (1994): *Adaptive pole placement without excitation probing signals.* — IEEE Trans. Automat. Contr., Vol.39, No.1, pp.47–59.

Ljung L. (1987): *System Identification: Theory for Users.* — Reading, MA: Addison-Wesley.

Nikolaou M. (1998): *Vision 2020: Workshop on Identification and Adaptive Control.* — Computers and Chemical Engineering, (in press).

Prett D.M. and Garcia C.E. (1988): *Fundamental Process Control.* — Reading, MA: Butterworths.

Shouche M., Genceli H. and Nikolaou M. (1998): *Simultaneous constrained model predictive control and identification of DARX processes.* — Automatica, Vol.34, No.12, pp.1521–1530..

Schwarm A.T, Eker S.A and Nikolaou M. (1998): *Model predictive control and identification: A new approach to closed-loop identification and adaptive control.* — Proc. Focapo Conf., Snowbird, Utah.

Söderström T. and Stoica P. (1989): *System Identification.* — Englewood Cliffs, NJ: Prentice-Hall.

Vandenberghe L. and Boyd S. (1996): *Semidefinite Programming.* — SIAM Review, Vol.38, No.1, pp.49–95.

Ydstie B.E. (1997): *Certainty equivalence adaptive control: What's new in the gap.* — Proc. 5-th Int. Conf. *Chemical Process Control*, AIChE Symposium Series, Vol.93, pp.9–23.

Ziegler J.G. and Nichols N.B. (1942): *Optimum settings for automatic controllers.* — Trans. ASME, Vol.64, No.7, pp.759–768.